

# Health Data Ecosystem: Readiness for Genomic Data Sharing

Mark Hoffman, Ph.D.

Vice President Cerner Life Sciences  
Solutions

# Topics

1. Generate high quality data during patient care to facilitate data sharing and decision support
2. Data sharing ecosystem examples

## Role of HIT:

- Enable healthcare system, academic, government, consumer-controlled and other stakeholder data sharing objectives through software and services

# Electronic Health Record

- Key attributes:
  - Capture information during clinical process
  - Simplify data retrieval, queries and analysis
  - Automate some processes
  - Provide decision support capabilities to prevent errors
  - Provide efficiencies
  - Generates body of data that can be analyzed to provide new administrative, operational, clinical and scientific insights
  - Legally binding medical record (discoverable)
  - PHR can complement EHR but is distinct

# Old world

	Breakfast			Lunch			Dinner			Bedtime			Other			Notes
	Dose	Blood	Sugar	Dose	Blood	Sugar	Dose	Blood	Sugar	Dose	Blood	Sugar	Dose	Blood	Sugar	
Mon		109			117			122			115					
Tues		111			106					*	152					* MISSED EVENING WALK. START BACK TOMORROW!
Wed		126			121			131			120					
Thurs		113			128			179		*	241					* SICK w/ FLU. DRINKING DIET SODA. KETONES NEGATIVE.
Fri		159			147			136			130					FEELING BETTER TODAY
Sat		127						124		*	152		130	11pm		* EMMA JUICE MADE SUGAR GO UP.
Sun		119			120			* 146			130					* LUNCH AT PICNIC.

# New world

## Health Record



My life, my health.

[Print Record](#)

To view your health record, click a section.

- ▶ [Allergies](#)
- ▶ [Clinical Notes](#)
- ▼ [Clinical Results](#)

Show Last:

Condition Center ☐

1 - 8 of 8

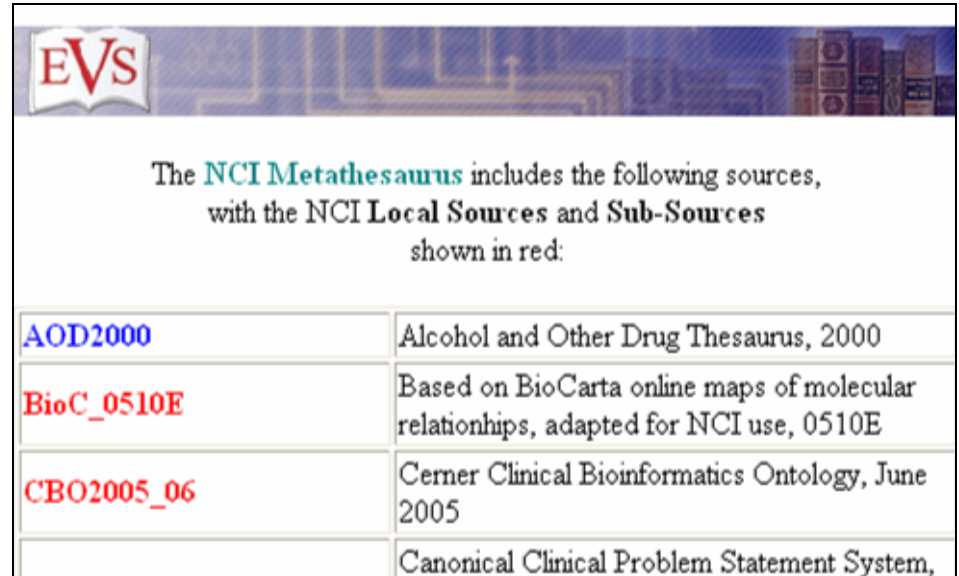
[Add New](#) | [Print](#)

Test Name	Result	Normal Range	Date/Time Performed	Source
<a href="#">Hemoglobin A1c</a>	7.1	7.0 - 7.5	10/02/2006 9:00 AM	Personal
<a href="#">Hemoglobin A1c</a>	7.2	7.0 - 7.5	9/28/2006 11:00 AM	Personal
<a href="#">Hemoglobin A1c</a>	7.3	7.0 - 7.5	9/01/2006 10:00 AM	Personal
<a href="#">Hemoglobin A1c</a>	7.2	7.0 - 7.5	8/15/2006 9:23 AM	Personal
<a href="#">Hemoglobin A1c</a>	7.0	7.0 - 7.5	8/10/2006	Personal
<a href="#">Hemoglobin A1c</a>	7.2	7.0 - 7.5	8/07/2006	Personal
<a href="#">Hemoglobin A1c</a>	7.7	7.0 - 7.5	7/25/2006	Personal
<a href="#">Microalbumin (Urine)</a>	.2	.001 - .4	7/03/2006 10:00 AM	Personal

- ▶ [Health Issues](#)
- ▶ [Immunizations](#)
- ▶ [Medications](#)
- ▶ [Surgeries/Procedures](#)

# The Clinical Bioinformatics Ontology (CBO)

- Scope:
  - Molecular pathology
  - Clinical Genetics
  - Infectious disease
  - Cytogenetics
  - Cell surface markers
- Current state (Jun 2009):
  - 12,510 Concepts
  - 29,355 Relationships
  - 761 Genes
  - 1,239 Nucleotide variants
  - 7,535 Terms
  - 6,095 Facets
- Open content approach
- Available from: [www.clinbioinformatics.org](http://www.clinbioinformatics.org)
- Included in the NCI Metathesaurus



The screenshot shows the NCI Metathesaurus interface. At the top is the EVS logo. Below it, text states: "The NCI Metathesaurus includes the following sources, with the NCI Local Sources and Sub-Sources shown in red:". A table follows, listing sources. The first source, AOD2000, is in blue. The subsequent three sources, BioC\_0510E, CBO2005\_06, and Canonical Clinical Problem Statement System, are in red.

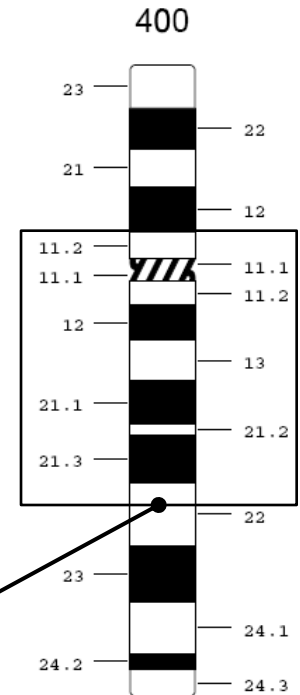
AOD2000	Alcohol and Other Drug Thesaurus, 2000
BioC_0510E	Based on BioCarta online maps of molecular relationships, adapted for NCI use, 0510E
CBO2005_06	Cerner Clinical Bioinformatics Ontology, June 2005
	Canonical Clinical Problem Statement System,

# Generate and codify complex discrete data during patient care

```
46,XY,add(1)(p36.1),der(8)?(qter->q22::p11.2->qter)[7]/47,XX,+2,der(8)?(qter->q22::p11.2->qter)[2]/47,XX,+8,der(8)?(qter->q22::p11.2->qter)[2]/46,XX[6]
```



Descriptor	Concept	Questionable
Sex Determinant	XY	<input type="checkbox"/>
Additional Material	400.1p36.1	<input type="checkbox"/>
Derivative Chromosome	Chromosome 8	<input type="checkbox"/>
Duplication	400.8p11.2	<input checked="" type="checkbox"/>
Duplication	400.8p11.1	<input checked="" type="checkbox"/>
Duplication	400.8p10	<input checked="" type="checkbox"/>
Duplication	400.8q10	<input checked="" type="checkbox"/>
Duplication	400.8q11.1	<input checked="" type="checkbox"/>
Duplication	400.8q11.2	<input checked="" type="checkbox"/>
Duplication	400.8q12	<input checked="" type="checkbox"/>
Duplication	400.8q13	<input checked="" type="checkbox"/>
Duplication	400.8q21.1	<input checked="" type="checkbox"/>
Duplication	400.8q21.2	<input checked="" type="checkbox"/>
Duplication	400.8q21.3	<input checked="" type="checkbox"/>



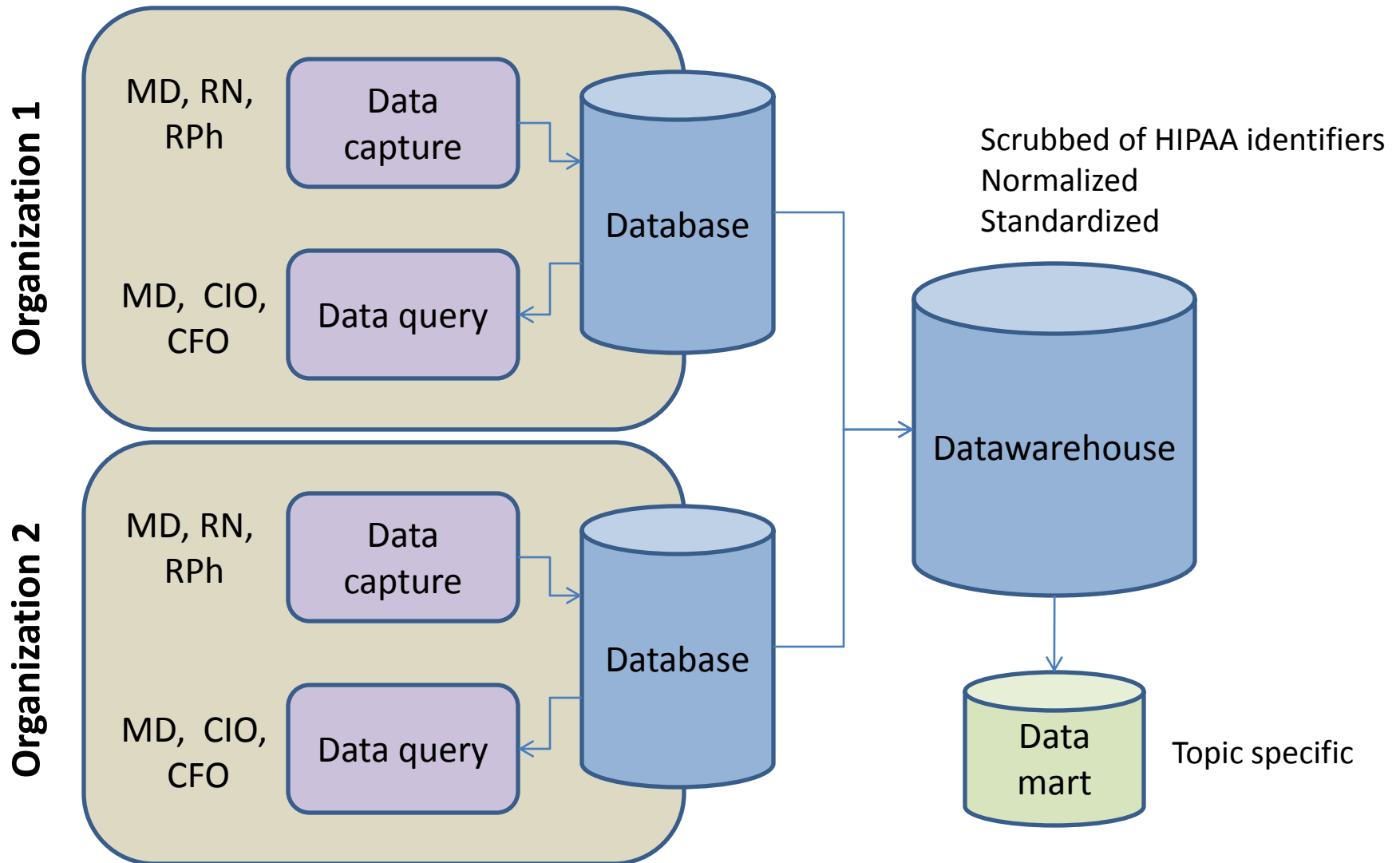
Implicit information  
would be missed by  
pure string search

# Representative data sharing models

- Centralized data warehouse
- Distributed queries using common technology platform
- Project-based data warehouse
- Consent based web system
- Social media

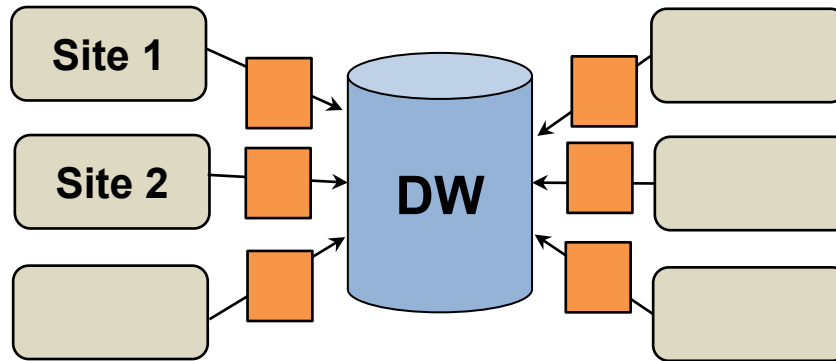


# Components of data sharing ecosystem

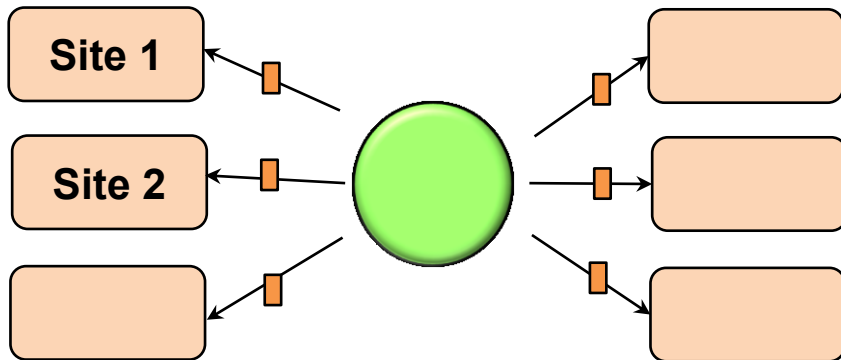


# Data warehouse compared to distributed model

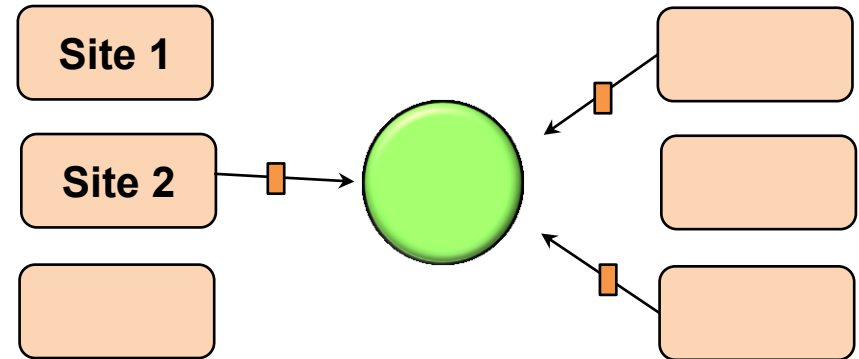
Data  
Warehouse



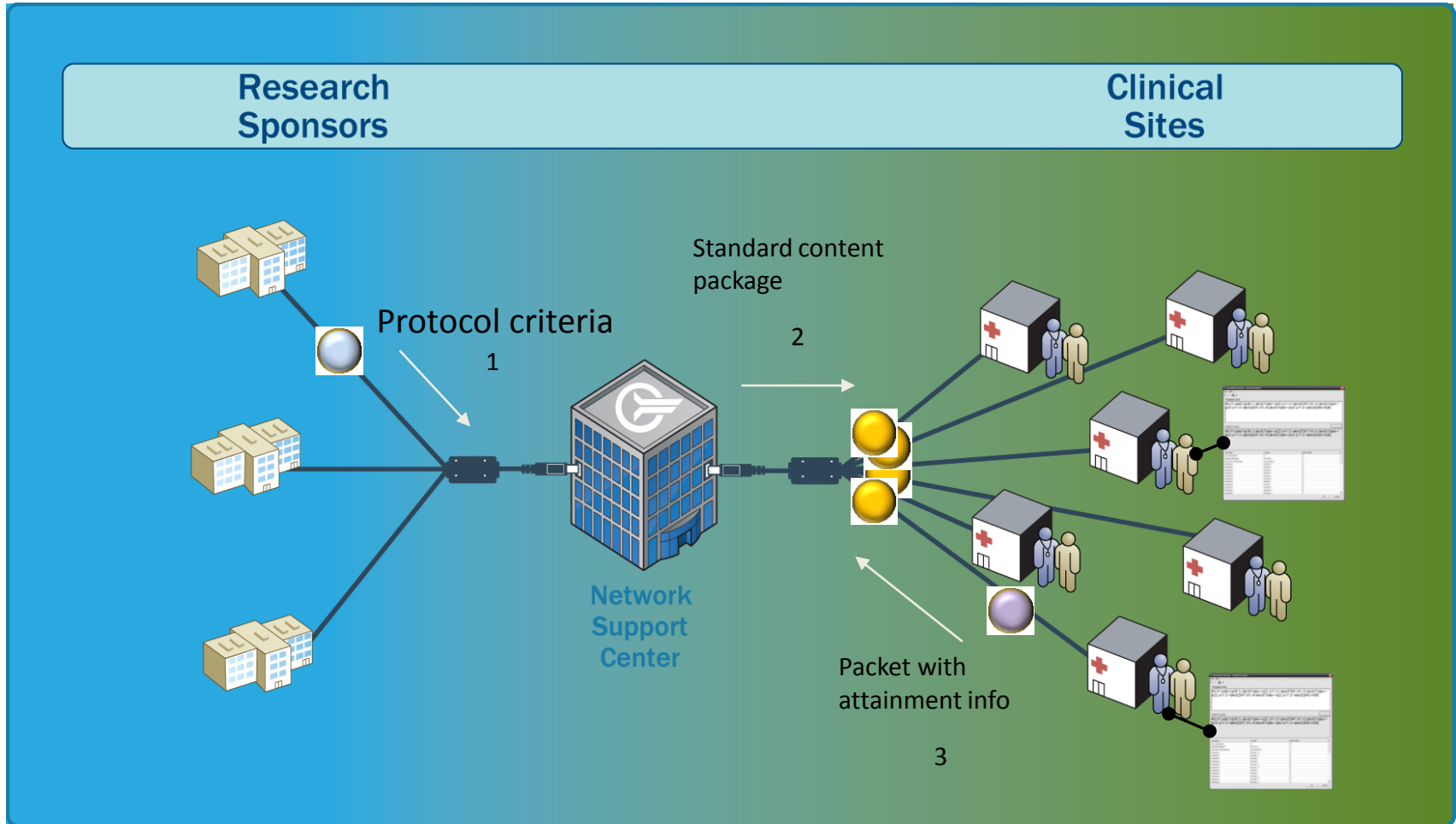
Send Queries



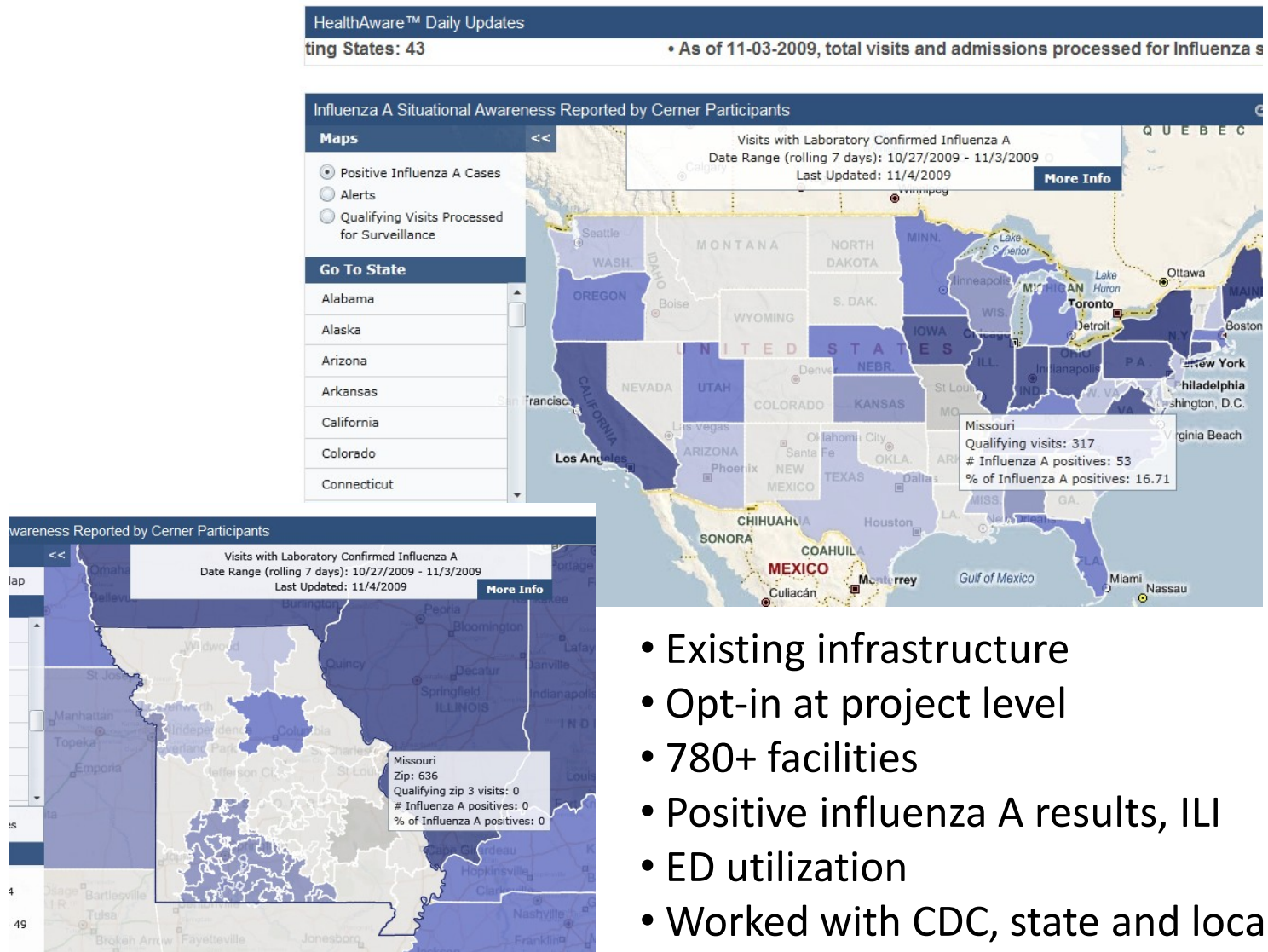
Receive attainment responses



# Research Network

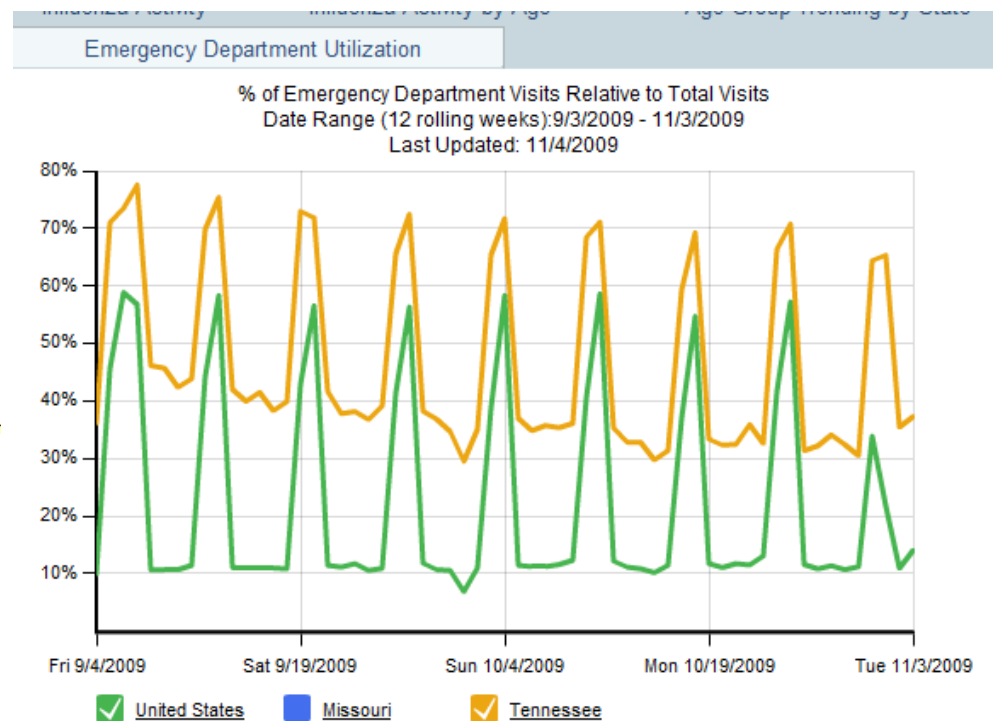
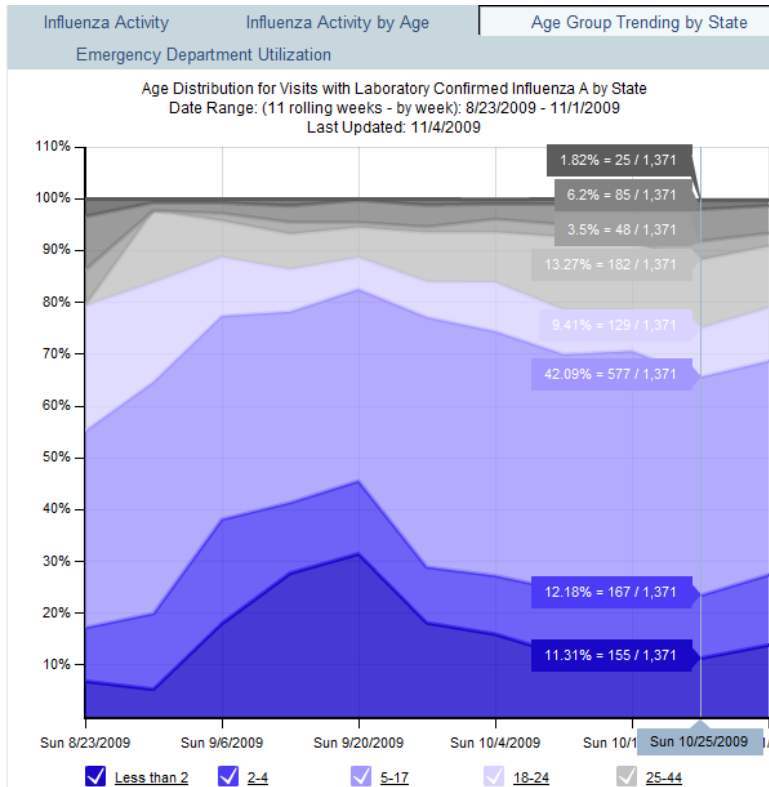


# Population monitoring: Influenza Initiative



- Existing infrastructure
- Opt-in at project level
- 780+ facilities
- Positive influenza A results, ILI
- ED utilization
- Worked with CDC, state and local public health

# Daily dashboard: A new population data paradigm

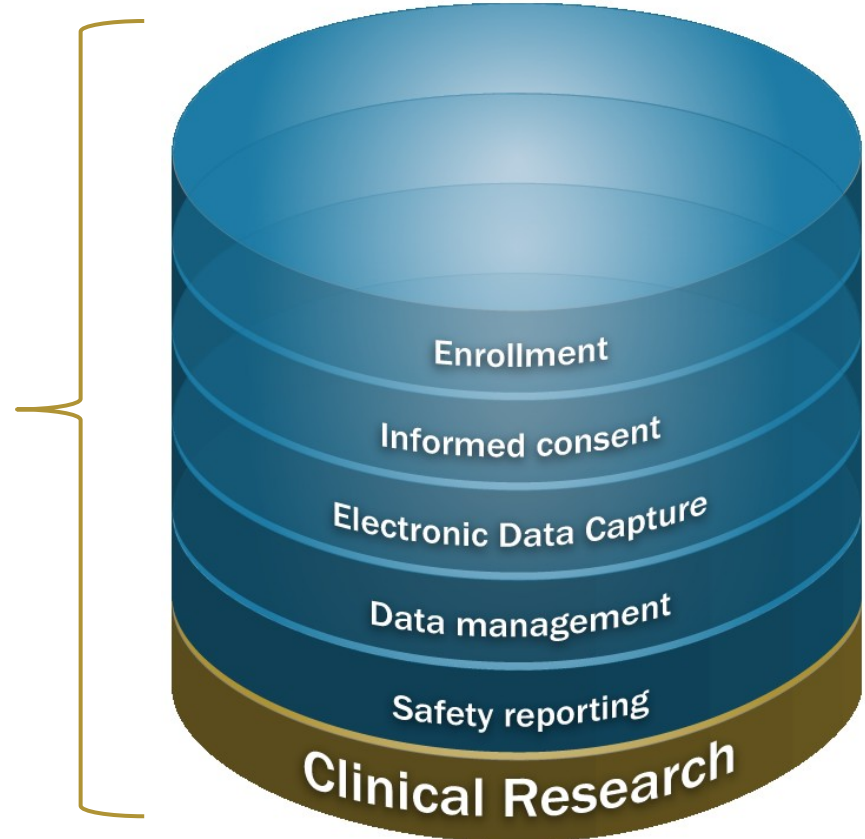
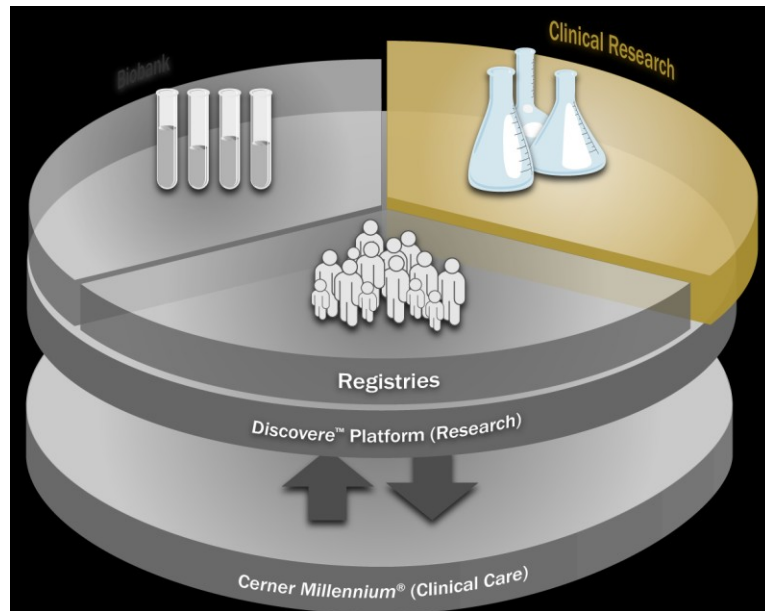


Tennessee

# Prospective Research Data: Consumer Involvement

## Discovere<sup>TM</sup>

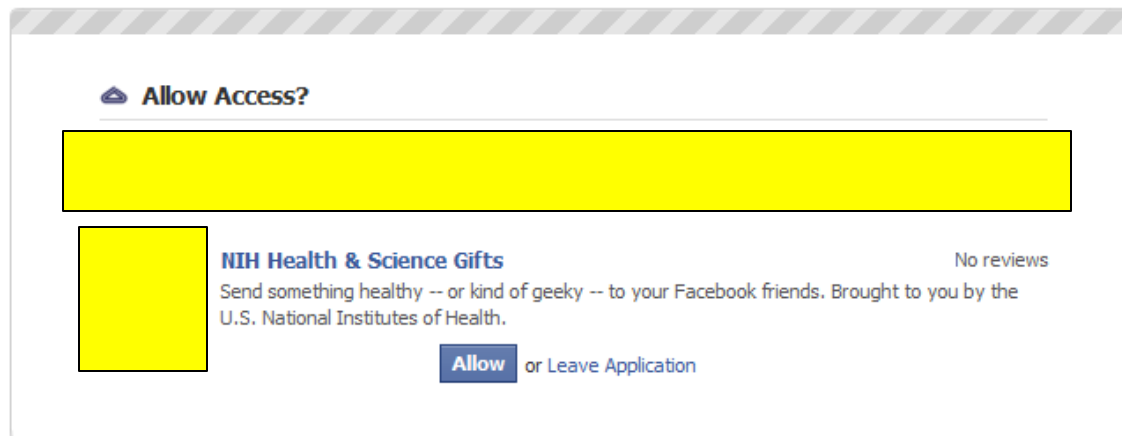
- Study Management
- Reporting Registries
- Biobanking
- Web based



# Population-based Discovery to Personalized Medicine

- What is the clinical impact of failure to appropriately use genotype results?
- **Approach:** Retrospective analysis of clinician behavior using HIV Outpatient Study data collection process
  - Multi-site CDC sponsored study (10 clinics in 7 cities), Cerner provides proj. mgmt and data mgmt
  - 8,000+ patients
    - 250,000 visits
    - Longitudinal data (fully consented)
- Evaluate use of HIV genotype information for antiretroviral selection
  - RT mutation for **D30N (n=67)** and **K103N (n=347)**
- Medication selection after genotype result
  - **441** patients with resistant genotype
  - **239** were receiving antiretroviral to which their HIV strain was resistant at the time of the test result
  - **42 (17.6%)** were still receiving AR with documented resistance 6 months after genotype
  - **59** - Initiated contraindicated AR after genotype results
- Clinical impact:
  - Fewer patients reached undetectable viral load or had drop in viral load ( $p < 0.05$ )
  - Drop in CD4 levels (not statistically significant)
- Explanations:
  - **36% acknowledged that they continued in error**
  - 22% limited options (“salvage” therapy)
  - 19% adherence issues, clinical trial or patient preference
  - **9 cases – could not locate records (partly reflecting reliance on paper reports)**

# Social Media – New dynamics, new issues for data sharing



ond There is a NEW Retinoblastoma Parental Support Network can join, meet other parents of RB kids, vent, laugh, and share r RB kids!!





# Topics for consideration

- Large aggregate data warehouse model:
  - Comprehensive nature allows rapid iterative analysis
  - High cost to operate and maintain
  - Requires transfer of data to central warehouse
  - Can support strategic analyses across non-affiliated organizations that would otherwise not be possible
  - Statistical power for Comparative Effectiveness Research
- Distributed models:
  - Limit data transfer
  - Summarization occurs locally at contributor site
  - Do not allow rapid iterative analysis
- Social Media
  - Much can be inferred by scanning affiliations, profiles and blogs
  - “I agree to share my data AND MY FRIENDS data”.
- Role of HIT:
  - Standardize and structure information at point of care
  - Provide software and services to enable data exchange
    - Participation is guided by IRBs and other policy decision makers among participating organizations
    - Deep expertise in clinical workflow, processes and data architecture